

# Flow Control for Cost-Efficient Peer-to-Peer Streaming

Dan-Cristian Tomozei  
Thomson

Paris Research Lab  
Email: dan-cristian.tomozei@thomson.net

Laurent Massoulié  
Thomson

Paris Research Lab  
Email: laurent.massoulie@thomson.net

**Abstract**—In this paper we address the issue of network cost efficiency for live streaming peer-to-peer systems. We formalize this as an optimization problem, which features a generic cost function. The latter is appropriate to capture not only ISP-specific link weights, but also non-linear, congestion-dependent costs. Our main contribution is the introduction of the *Implicit-Primal-Dual* scheme for flow control in live streaming peer-to-peer systems. It is fully distributed in that it relies only on local state variable exchanges. Moreover, we show that at a fluid scale, combined with random linear network coding, it admits the cost optimal operating point as a fixed point. We also prove asymptotic boundedness of fluid trajectories for particular cost functions. We finally show via experiments that these optimality properties are resilient to operational constraints such as finite generation size and finite field size.

## I. INTRODUCTION

Peer-to-peer systems for live streaming, such as TVAnts and Sopcast, have recently met considerable success. As testimony for this, PPLive, which enjoys over 100 million users, has obtained in 2008 a license from the Chinese government to broadcast media content. Despite this growth and evolution to mainstream usage (as opposed to underground, illegal usage), peer-to-peer systems have recently been facing strong criticisms from ISPs.

Indeed, such systems generate massive amounts of traffic on sensitive physical links, both in the access and in the core networks. One solution to the problem, used by ISPs, consists in throttling peer-to-peer traffic at such critical parts of the network, using hardware dedicated to this task.

Another approach, under discussion at the P4P working group, consists in letting the ISPs provide information to the peer-to-peer system trackers about their network costs. The hope is that the peer-to-peer system will then be able to adapt its traffic flows to reduce congestion at hot spots within the network, without sacrificing user experience. This strategy is more appealing than the previous one, in that it could lead to a “win-win” outcome for network operators and end users, rather than an arms race.

It is however challenging to implement. Most prominently, it requires from the peer-to-peer system design the ability to exploit information about the underlying network costs and regulate data flows accordingly between large numbers of peers. The scale of such systems clearly mandates lightweight, distributed solutions. Our aim in the present paper is

to propose such solutions.

To address the problem, we take the following approach. We first propose a formulation in terms of an optimization problem, featuring a cost function to be minimized, given some feasibility conditions related to “cut constraints”. Inspired by a primal-dual algorithmic solution to this problem, we develop our so-called *Implicit-Primal-Dual* (IPD) control scheme. The latter is fully distributed, in that neighbor peers adjust their data rates based only on local information exchanges. It relies on the use of specific quantities of backlogged information as implicit “multipliers” for the cut constraints at hand.

We then illustrate how it might fall short of achieving a global optimum through a fluid scale analysis of a particular scheduling strategy for packet transmission among peers. This motivates its combination with Random Linear Network Coding instead of pure store and forward for data exchanges.

Our main result is then a proof that under such Random Linear Coding, IPD admits the global system optimum as a stationary point, in a fluid regime. A byproduct is a modeling and fluid level description of networks under Random Linear Coding. We complement this result by a proof of asymptotic stability when the cost function is of so-called *separable* type. We finally provide experimental evidence that IPD, combined with Random Linear Coding, indeed achieves close to optimal performance even in non-fluid regimes. In particular we show that performance is resilient to changes in parameters essential in network coding, namely generation and field size.

The outline of the paper is as follows: In Section III we give the problem statement and introduce Implicit-Primal-Dual. In Section IV we analyze its performance when combined with the Random Useful (RU) packet selection strategy. In Section V we analyze IPD in combination with Random Linear Coding (RLC). We prove that over coded packet networks the global optimum is an equilibrium point for IPD. In Section VI we provide numerical results. We conclude in Section VII.

## II. RELATED WORK

In [1] the authors wish to maximize a utility function in a similar setting to the one described in this paper. They present a solution involving packing a polynomial number of depth-1 and -2 Steiner trees for multisource multicast. The scalability of their approach is limited as it requires all peers to maintain data exchanges with all other peers.

In [2] the authors consider the slightly different problem of multicast over coded wired and wireless networks with no required rate at the receivers. A message passing scheme is used to implement a maximum flow algorithm. The scalability of this approach is limited due to the fact that messages relative to each receiver need to be passed throughout the network. Thus the number of state variables to be maintained at each node scales linearly with the number of receivers.

In [3], the authors propose the so-called UESSM selfish routing algorithm in a similar setting to [2]. It is shown that UESSM converges to the min-cost flow allocation for a particular type of cost function. This approach has the same scalability limitations as [2].

To our knowledge, there have not been any proposed scalable distributed algorithms achieving minimum cost multicast.

### III. EXPLICIT AND IMPLICIT PRIMAL-DUAL

Consider a peer-to-peer network described by a graph  $\mathcal{G} = (\{s\} \cup \mathcal{V}, \mathcal{L})$  with an achievable capacity region  $\mathcal{C} \subset \mathbb{R}_+^{|\mathcal{L}|}$  (i.e. only rate allocations  $\vec{c} \in \mathcal{C}$  are feasible). The source  $s$  wishes to convey information (a stream of bytes) at a rate  $\lambda$  to the set of receivers  $\mathcal{V}$ . We define  $c_{ij}$  as the rate at which information is transferred along edge  $(i, j) \in \mathcal{L}$ . Also consider an increasing strictly convex differentiable function  $\Gamma(\vec{c})$  describing the global cost incurred from transferring information on edges at rates given by  $\vec{c} \in \mathbb{R}_+^{|\mathcal{L}|}$ . We wish to minimize cost while still achieving a transfer rate of  $\lambda$  at each of the receivers in  $\mathcal{V}$ .

More formally, denote by  $\mathcal{T}_s = \{\{s\} \cup S' : S' \subsetneq \mathcal{V}\}$  the set of non-trivial cuts containing the source node and by  $c(S, \vec{S})$  for  $S \in \mathcal{T}_s$  the transmission rate through cut  $S$ , namely  $\sum_{i,j:i|Sj} c_{ij}$ . The notation “ $i|Sj$ ” signifies  $i \in S$ ,  $j \notin S$  and  $(i, j) \in \mathcal{L}$  (i.e. cut  $S$  separates node  $i$  from node  $j$  or, equivalently, link  $(i, j)$  crosses cut  $S$ ) and will be used frequently in the remainder of the paper. We state the following optimization problem which describes our goal:

$$\text{minimize } \Gamma(\vec{c}) \text{ over } \vec{c} \quad (1)$$

$$\text{subject to } c(S, \vec{S}) \geq \lambda, \quad S \in \mathcal{T}_s. \quad (2)$$

The cut constraint (2) is clearly necessary. It is also sufficient when network coding is allowed, in view of the founding result in network coding, identifying feasible multicast rates with the “min-min-cut” [4]. In the present context where all nodes are receivers, the feasible broadcast rate is again given by the min-min-cut condition, as follows from Edmonds’ theorem [5].

We suppose the capacity constraints are included in the cost function. Thus, a further condition the cost function must satisfy is  $\Gamma(\vec{c}) = +\infty$  for  $\vec{c} \notin \mathcal{C}$ . Denote the marginal cost of a link  $\ell$  by  $p_\ell(\vec{c}) = \frac{\partial \Gamma}{\partial c_\ell}(\vec{c})$ . We will also refer to this marginal cost as the *price* of link  $\ell$ .

For example take a network with  $R$  physical links and  $L$  overlay links. The underlay routing can be expressed as a routing matrix  $H \in \{0, 1\}^{R \times L}$ , for which

$$H_{r\ell} = \begin{cases} 1 & \text{if overlay link } \ell \text{ uses physical link } r, \\ 0 & \text{otherwise.} \end{cases}$$

Denote by  $\sigma_r(\vec{c}) = C_r - \langle \vec{H}_r, \vec{c} \rangle$  the spare capacity on link  $r$ . A possible choice of cost function is network congestion, i.e. setting link prices equal to the observed packet delay. Such prices can be modeled as

$$p_\ell(\vec{c}) = \sum_{r: H_{r\ell}=1} \frac{1}{\sigma_r(\vec{c})}. \quad (3)$$

Furthermore, ISPs wishing to avoid overloading sensitive links can define physical link weights  $w_r$  which increase additively the price of an overlay link:

$$p_\ell^{ISP}(\vec{c}) = \sum_{r: H_{r\ell}=1} \left( w_r + \frac{1}{\sigma_r(\vec{c})} \right).$$

The corresponding global cost function writes as:

$$\Gamma_{ISP}(\vec{c}) = \sum_{r=1}^R w_r \langle \vec{H}_r, \vec{c} \rangle - \sum_{r=1}^R \log \frac{\sigma_r(\vec{c})}{C_r}.$$

Let us now characterize the optimum rates minimizing some generic  $\Gamma$ . For problem (1–2) the Lagrangian is:

$$\mathcal{L}(\vec{c}, \vec{\mu}) = \Gamma(\vec{c}) + \sum_{S \in \mathcal{T}_s} \mu_S (\lambda - c(S, \vec{S})), \quad (4)$$

where the  $\{\mu_S\}_{S \in \mathcal{T}_s}$  are Lagrange multipliers associated with the cut constraints (2). Taking the partial derivative with respect to  $c_{ij}$  for some  $(i, j) \in \mathcal{L}$ , we obtain that the optimum values  $(\vec{c}^*, \vec{\mu}^*)$  satisfy

$$p_{ij}(\vec{c}^*) = \sum_{S: i|Sj} \mu_S^*, \quad (5)$$

where the Lagrange multipliers  $\mu_S^*$  satisfy the KKT conditions:

$$\mu_S^* \geq 0, \quad (6)$$

$$\mu_S^* (\lambda - c^*(S, \vec{S})) = 0. \quad (7)$$

We are interested in solving this problem in a fully distributed fashion. A classic solution for solving constrained optimization problems is the “Primal-Dual” method. This approach consists of letting the primal and dual variables evolve simultaneously towards the saddle point of (4). Srikant [6] applies this method to the field of optimal resource allocation and congestion control algorithms. In the case of our problem, the Primal-Dual algorithm takes the form of the following autonomous differential system:

$$\dot{c}_{ij}(t) \propto \left[ \sum_{S: i|Sj} \mu_S(t) - p_{ij}(\vec{c}(t)) \right]_{c_{ij}(t)}^+, \quad (i, j) \in \mathcal{L} \quad (8)$$

$$\dot{\mu}_S(t) \propto \lambda - \sum_{i|Sj} c_{ij}(t), \quad S \in \mathcal{T}_s. \quad (9)$$

Above we used a convenient notation:

$$(a)_b^+ = \begin{cases} a & \text{if } b > 0, \\ \max\{a, 0\} & \text{if } b = 0. \end{cases}$$

Proofs of asymptotic stability and optimality of the Primal-Dual algorithm (8-9) can be found in [6](pp. 32–33). Note that

a violation of the constraint (2) for some cut  $S$  will induce an increase in the associated quantity  $\mu_S$  by (9), which will in turn induce an increase in the transmission rate through cut  $S$  by (8). This rate increase will continue until the said cut constraint is satisfied.

Despite the nice convergence and stability properties of this scheme, it is not easily implementable in practice. Since there exists an exponential number of dual variables, a direct implementation of the Primal-Dual algorithm would require maintenance of an exponential number of state variables, which is prohibitive.

In most peer-to-peer protocols data is divided into small pieces called *chunks*. A chunk constitutes the basic data transfer unit in such a network. For the rest of this paper we shall refer to such chunks as “packets”. Assume node  $i$  has a set  $P_i(t)$  of packets at time  $t$ . For all sets of nodes  $S$ , denote by  $Z_S(t)$  the number of packets which are found at every node in  $S$  and nowhere outside of  $S$  at time  $t$  (i.e. the backlog at  $S$ ). More formally,  $Z_S(t) = |\bigcap_{i \in S} P_i(t) \setminus \bigcup_{j \in S^c} P_j(t)|$ .

Define  $X_{+i-j}(t) = |P_i(t) \setminus P_j(t)|$ . These quantities describe the number of packets present at node  $i$  and absent at node  $j$  at time  $t$ . By elementary methods it can be shown that

$$X_{+i-j}(t) = \sum_{S: i|_S j} Z_S(t). \quad (10)$$

We notice a strong similarity between (10) and the contribution  $\sum_{S: i|_S j} \mu_S(t)$  of the dual variables to the rate adaptation rule (8) in the Primal-Dual algorithm. Furthermore, the  $X_{+i-j}$  can be computed *locally* to link  $(i, j)$  without specific knowledge of quantities  $Z_S$ . This observation makes the backlog sizes  $Z_S$  excellent candidates for approximating the dual variables  $\mu_S$  and eliminates the need for transferring exponentially many control messages for computing updated rate values. Furthermore, backlog sizes evolve *implicitly* and their specific evolution depends on the employed packet scheduling strategy. We do not need to keep track of them in an explicit fashion.

Motivated by the previous remarks, we propose *Implicit-Primal-Dual* (IPD), a distributed rate control algorithm, for which we use a continuous rate adaptation method, the *Implicit-Primal* rule (as opposed to the *Explicit-Primal* (8)):

$$\dot{c}_{ij}(t) = \kappa [\alpha X_{+i-j}(t) - p_{ij}(\vec{c}(t))]_{c_{ij}(t)}^+, \quad (11)$$

for all  $(i, j) \in \mathcal{L}$ . Here  $\kappa$  is a gain factor and  $\alpha$  is a conversion factor (it can be considered as price per useful packet). Further intuition behind this rule is as follows: the more a node  $i$  has to offer to its neighbor  $j$ , the more it will increase the rate  $c_{ij}$ ; at the same time, the higher the price of the link, the lower the rate  $c_{ij}$  will become.

As we previously stressed, the *Implicit-Dual* equations describe the backlog size evolution and follow implicitly from the packet transfer strategy. In the next Sections we consider a classic packet forwarding network implementing the Random Useful (RU) scheme and a coded packet network implementing Random Linear Coding (RLC). We use a fluid scaling of the system to deduce the governing Implicit Dual rule for both cases. We find that IPD deviates from the optimum operation

point under RU due to redundant transmissions over congested links. We show that RLC solves this problem and that the global optimum is a stationary point for IPD in this setting.

#### IV. A FIRST APPROACH

Consider the case in which the source generates a sequence of packets it wishes to deliver to the receivers in  $\mathcal{V}$  at rate  $\lambda$ . We use the notations  $P_i, Z_S, X_{+i-j}$  introduced in the previous Section III. Nodes are assumed to have the knowledge of which packets are present at their one-hop neighbors. Consider the following transmission strategy: For each link  $(i, j) \in \mathcal{L}$ , with rate  $c_{ij}(t)$  node  $i$  picks a packet uniformly at random from the set  $P_i(t) \setminus P_j(t)$  of packets useful for  $j$  and sends it along the link. Such a transmission occurs if and only if  $X_{+i-j}(t) = |P_i(t) \setminus P_j(t)| > 0$ . This strategy is called the *Random Useful* strategy. In [7] the authors prove that Random Useful (RU) is rate-optimal for a fixed rate allocation  $\vec{c}_0$ . That is, when using RU if the value of the min-min-cut of the edge-capacitated graph  $(\mathcal{G}, \vec{c}_0)$  with source node  $s$  is greater than  $\lambda$ , then rate  $\lambda$  is attained at each receiver. Starting from this rate optimal strategy, we wish to characterize the behavior of the proposed IPD rate control scheme (11).

We model the system as follows: We assume for simplicity that packets are generated at the source at instants of a Poisson process of rate  $\lambda$ . We also assume that packet transfers on links  $\ell \in \mathcal{L}$  are triggered at instants of Poisson processes of time-varying rates  $c_\ell(t)$ . The process pair  $((c_\ell(t))_{\ell \in \mathcal{L}}, (Z_S(t))_{S \in \mathcal{T}_s})$  is jointly Markovian with a continuous and a discrete component. Packet generation at the source increments  $Z_{\{s\}}$  (at rate  $\lambda$ ). Along any edge  $(i, j) \in \mathcal{L}$  there are  $X_{+i-j}$  packets available for transfer from node  $i$  to node  $j$ . From (10) it follows that for every packet selected for transfer along edge  $(i, j)$  (according to the RU strategy) there exists a unique set  $S \in \mathcal{T}_s$  with  $i|_S j$ , such that the selected packet belongs to the backlog at  $S$ . Such a packet transfer will decrement the corresponding  $Z_S$  and increment  $Z_{S \cup j}$  simultaneously, and will occur with probability  $\frac{Z_S}{X_{+i-j}}$ . Moreover, since the process dictating transfers is Poisson with rate  $c_{ij}$ , the simultaneous updates for  $Z_S$  and  $Z_{S \cup j}$  will be performed at rate  $c_{ij} \frac{Z_S}{X_{+i-j}}$ , provided that  $X_{+i-j} > 0$ .

In what follows we describe the behavior of the system at a fluid time scale. Based on this scaling, we show on an example that IPD is suboptimal. We analyze the reason for which IPD deviates from the optimum point and we propose a coding-based approach.

##### A. Fluid Limits

We examine a scaling of (11) for which both  $\alpha$  and  $\kappa$  are small. Consider the sequence of processes  $(c^N, Z^N)_N = ((c_\ell^N(t))_{\ell \in \mathcal{L}}, (Z_S^N(t))_{S \in \mathcal{T}_s})_N$  evolving according to dynamics introduced above, such that the  $N^{\text{th}}$  process follows IPD rule (11) with  $\alpha = \frac{1}{N}$  and  $\kappa = \frac{g}{N}$ , for some positive constant gain  $g$ . Assume initial conditions  $(c^N(0), Z^N(0))_N$  such that

$$\lim_{N \rightarrow \infty} c_\ell^N(0) = c_\ell, \quad \lim_{N \rightarrow \infty} \frac{1}{N} Z_S^N(0) = z_S.$$

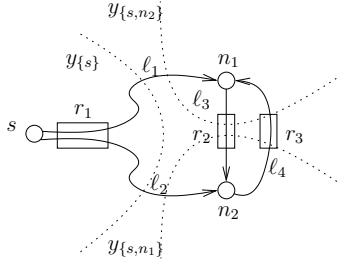


Fig. 1: A counterexample illustrating the suboptimality of the Random Useful scheme.

We define  $S(c, y)$ , for  $c \in \mathbb{R}_+^{|\mathcal{L}|}$  and  $y \in \mathbb{R}_+^{|\mathcal{T}_s|}$ , the set of fluid trajectories with initial conditions  $c_\ell(0) = c_\ell$  and  $y_S(0) = y_S := \sum_{T \subset S} z_T$  as absolutely continuous functions  $c_\ell, y_S : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , such that:

$$\frac{d}{dt} c_{ij}(t) = g[y_{+i-j}(t) - p_{ij}(\vec{c}(t))], \quad (12)$$

$$\frac{d}{dt} y_{\{s\}}(t) = \lambda - \sum_{j \neq s} c_{sj}(t) \frac{y_{\{s\}}(t)}{y_{+s-j}(t)}, \quad (13)$$

$$\frac{d}{dt} y_S(t) = \lambda - \sum_{i,j: i \in S, j \notin S} c_{ij}(t) \frac{y_S(t) - y_{S-i}(t)}{y_{+i-j}(t)}, \quad (14)$$

almost everywhere, for all links  $(i, j) \in \mathcal{L}$  and all cuts  $S \in \mathcal{T}_s \setminus \{s\}$ , where  $y_{+i-j}(t) := y_{\overline{\{i,j\}}}(t) - y_{\{i,j\}}(t)$ .

Introduce the rescaled processes  $\hat{c}^N(t) = c^N(Nt)$  and  $Y_S^N(t) = \frac{\sum_{T \subset S} Z_T^N(Nt)}{N}$  for all cuts  $S \in \mathcal{T}_s$ . Similarly to [7] it can be shown that the processes  $(\hat{c}^N, Y^N)_N$  converge in probability for the topology of uniform convergence on finite intervals to the set of fluid trajectories  $S(c, y)$  as  $N \rightarrow \infty$ .

### B. Suboptimality

Let us show that, in the case of Random Useful, the global optimum is not always a fixed point of IPD. We present a simple counterexample. Take the network in Figure 1 and consider the following cost function:

$$\Gamma(\vec{c}) = -\log(\lambda + 2\epsilon - c_{s1} - c_{s2}) - \log(\lambda - c_{12}) - \log(\lambda - c_{21}).$$

This means that we suppose a maximum rate of  $\lambda + 2\epsilon$  on physical link  $r_1$  and a maximum rate of  $\lambda$  on physical links  $r_2$  and  $r_3$ . Assume  $\epsilon < \frac{\lambda}{4}$ . Computing the optimum yields

$$\begin{aligned} c_{s1}^* &= c_{s2}^* = c_{12}^* = c_{21}^* = \frac{\lambda}{2}, \\ \mu_{s1}^* &= \mu_{s2}^* = \frac{2}{\lambda}, \mu_s^* = \frac{1}{2\epsilon} - \frac{2}{\lambda}. \end{aligned}$$

We search a symmetrical equilibrium point of IPD such that  $c_{s1} = c_{s2} = \delta$  and  $c_{12} = c_{21} = \lambda - \delta$ , where  $0 < \delta < \lambda$ . Should an equilibrium point be asymmetrical, clearly it would deviate from the symmetrical optimal point. Then (13)–(12) give the following equations (and their symmetric versions which we do not write here) from which we wish to determine  $\delta$  at

the equilibrium:

$$\begin{aligned} \dot{c}_{s1} &= y_{s2} - \frac{1}{\lambda + 2\epsilon - c_{s1} - c_{s2}}, \\ \dot{c}_{12} &= y_{s1} - y_s - \frac{1}{\lambda - c_{12}}, \\ \dot{y}_s &= \lambda - c_{s1} \frac{y_s}{y_{s2}} - c_{s2} \frac{y_s}{y_{s1}}, \\ \dot{y}_{s1} &= \lambda - c_{s2} \frac{y_{s1}}{y_{s1}} - c_{12} \frac{y_{s1} - y_s}{y_{s1} - y_s}. \end{aligned}$$

We obtain that

$$\delta = \frac{\lambda}{2} + \frac{2\epsilon}{3}.$$

We have shown that the IPD scheme is suboptimal when Random Useful is used for the topology in Figure 1. The intuition behind suboptimal performance on this simple topology is as follows: The source will always send useful packets to both  $n_1$  and  $n_2$ . However it does not perform any coordination on which packets it sends to one or the other. Thus packet sets at  $n_1$  and at  $n_2$  are not “as different” as they could be. This causes a lower number of transmission opportunities for links  $\ell_3$  and  $\ell_4$ . In turn, this causes an increase in traffic on the congested physical link  $r_1$  and in global cost and thus a deviation from the global optimum. In the next section we solve this problem by using *Random Linear Network Coding*.

While we do not characterize formally the deviation from the global optimum, we observe by numerical evaluation provided in Section VI that IPD has “reasonably” good performance in the RU setting. As we previously stressed, it has a major advantage over a Primal-Dual scheme: its implementation only requires adding the local rate control rule to an already implemented Random Useful live streaming protocol.

## V. IPD WITH RANDOM LINEAR CODING

We now consider the combination of IPD with data transfers based on Random Linear Coding (RLC) [4] rather than RU packet forwarding. We first describe the system mechanisms and a corresponding stochastic model. We then characterize fluid limits for this model. Finally, we establish that for general convex cost functions, the optimal operating point is a stationary point of these fluid limits. We further show stability of these limits in the case of separable convex cost functions.

### A. System Mechanisms and Model

As in the previous section, the source generates packets, which we call *native* packets. However it sends out linear combinations of the native packets it has generated so far, rather than native packets. Receiver nodes accumulate linearly independent combinations of native packets. An incoming packet is kept only when it is deemed *innovative*, that is when it is linearly independent of those previously received. At each node, packets sent are linear combinations of the innovative packets available at this particular node. Nodes are able to decode a given set of  $K$  native packets as soon as they have received  $K$  linear combinations of these packets that are linearly independent.



Implementations of network coding for live-streaming rely on the so-called “generation” approach. There the stream is divided into “generations” of  $K$  consecutive native packets. Encoding and recoding operations are performed only among packets relative to the same generation. Users are then able to play content within a generation as soon as they have received  $K$  linearly independent combinations of packets within this generation. This approach has also been considered for Video-on-Demand applications in [8].

In the sequel we ignore the impact of finite generations, and assume instead that each packet transmitted is a random linear combination of *all* innovative packets previously received by the sending node. This assumption is made for the sake of analysis, as it greatly simplifies modeling of the system. The impact of finite generations on performance will be assessed experimentally in Section VI. We also assume that the linear coefficients used in the construction of an encoded packet are taken uniformly at random in the finite field  $\mathbb{F}$  over which linear operations are performed.

Let us now introduce some notation. The  $i^{th}$  native packet is associated with the  $i^{th}$  unit vector in  $\mathbb{F}^N$ . Any data packet is associated with the vector of coefficients in  $\mathbb{F}^N$  characterizing its combination in terms of native packets. For any user  $i$ , we denote by  $V_i$  the subspace of  $\mathbb{F}^N$  spanned by the packets it owns. Note that the vector space  $V_i$  evolves with time as new packets reach node  $i$ .

For any set  $S$  of nodes, we shall also denote by  $V_S$  the vector space spanned by the packets owned by all nodes in  $S$ :

$$V_S = \sum_{i \in S} V_i. \quad (15)$$

We further define the quantity  $X_S$  as the number of innovative packets that nodes in  $S$  could provide to nodes in  $\bar{S}$ , that can be written as

$$X_S = \dim(V_S) - \dim(V_{\bar{S}}). \quad (16)$$

Indeed, the validity of this expression can be readily checked by remarking that  $V_S \subset V_s$  for all sets  $S$ , and  $V_S = V_s$  when  $s \in S$ .

Finally, we let  $X_{i-j}$  denote the number of innovative packets that  $i$  could send to  $j$ , which can be written as

$$X_{i-j} := \dim(V_i + V_j) - \dim(V_j) = X_{\{j\}} - X_{\{i,j\}}. \quad (17)$$

In the present context, the Implicit-Primal-Dual adaptation rule of the transmission rates  $c_{ij}$  is again defined by (11), based on the above definition (17) of  $X_{i-j}$ .

To complete our formal model, we make the following statistical assumptions. Native packets are created at the source  $s$  at the instants of a Poisson process with rate  $\lambda > 0$ . For each link  $(i, j) \in \mathcal{L}$ , packets are transmitted from  $i$  to  $j$  at the instants of a Poisson process with time-varying rate  $c_{ij}$ . Such packets correspond to a uniform random linear combination of the innovative packets received by node  $i$  at the time of the transition.

Thus at rate  $\lambda$ , the vector space  $V_s$  increases its dimension by 1, and hence quantities  $X_S$  simultaneously increase by 1

for all  $S \in \mathcal{T}_s$ . Also, at rate  $c_{ij}$ , for sets  $S \in \mathcal{T}_s$  with  $i \in S$ ,  $j \notin S$ , quantities  $X_S$  decrease by 1 provided the random linear combination formed at  $i$  does not belong to  $V_{\bar{S}}$ .

With the above assumptions, the pair of processes

$$((c_\ell(t))_{\ell \in \mathcal{L}}, (V_i(t))_{i \in \{s\} \cup \mathcal{V}})$$

is jointly Markovian, with both a continuous and a discrete component.

Note that a uniform random selection of a vector in  $V_i$  will belong to  $V_{\bar{S}}$  with probability

$$r(S - i, S) := \frac{|V_i \cap V_{\bar{S}}|}{|V_i|} = \left( \frac{1}{|\mathbb{F}|} \right)^{X_S - X_{S-i}}. \quad (18)$$

The probability that such a transmission leads to a decrease in  $X_S$  thus reads  $1 - r(S - i, S)$ .

It should be noted that for two distinct sets  $S, T$  both containing  $i$  and not containing  $j$ , the occurrences of increase in  $X_S$  and  $X_T$  are not independent, their joint probability depending in an intricate fashion on the sizes of the spaces  $V_i, V_{\bar{S}}, V_{\bar{T}}$  and of their intersections. These correlations will however not play any role in the behavior of the system in the fluid limit, which we now characterize.

### B. System Dynamics in the Fluid Limit

We shall again consider the system in the limit where both the scale parameter  $\alpha$  and the gain parameter  $\kappa$  appearing in (11) are small. The limiting dynamics will be related to the fluid trajectories that we now define:

*Definition 1:* The functions  $c_{ij}, v_S : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , for  $(i, j) \in \mathcal{L}$ ,  $S \in \mathcal{T}_s$  are called fluid trajectories of the RLC streaming system if they are absolutely continuous, and admit the following derivatives almost everywhere:

$$\frac{d}{dt} c_{ij}(t) = g[v_{i-j}(t) - p_{ij}(\bar{c}(t))], \quad (i, j) \in \mathcal{L}, \quad (19)$$

$$\frac{d}{dt} v_S(t) = \begin{cases} \lambda - \sum_{i \in S, j \in \bar{S}} c_{ij}(t) + \sum_{i \in S, j \in \bar{S}} c_{ij}(t) a_{S-i, S}(t), & S \in \mathcal{T}_s \setminus \{s\}, \\ \lambda - \sum_{j \neq s} c_{sj}(t) + \sum_{j \neq s} c_{sj}(t) a_s(t), & S = \{s\}, \end{cases} \quad (20)$$

where  $g$  is some positive gain,  $v_{i-j}(t) := v_{\{j\}}(t) - v_{\{i,j\}}(t)$ , and the functions  $a_{S-i, S}(t), a_s(t)$  verify

$$\begin{aligned} a_{S-i, S}(t), a_s(t) &\geq 0, \\ a_{S-i, S}(t) [v_S(t) - v_{S-i}(t)] &\equiv 0, \\ a_s v_{\{s\}}(t) &= 0. \end{aligned} \quad (21)$$

In addition, the non-negative functions  $v_S$  are required to verify

$$S \subset T \Rightarrow v_S \leq v_T, \quad \forall S, T \in \mathcal{T}_s, \quad (22)$$

as well as the supermodularity condition:

$$v_S + v_T \leq v_{S \cup T} + v_{S \cap T}, \quad \forall S, T \in \mathcal{T}_s. \quad (23)$$

For any  $c \in \mathbb{R}_+^{|\mathcal{L}|}$  and  $v \in \mathbb{R}_+^{|\mathcal{T}_s|}$ , we shall further denote by  $S(c, v)$  the set of all such fluid trajectories with initial conditions  $c(0) = c, v(0) = v$ .

We then have the following:

*Theorem 1:* Consider a sequence of processes  $(c^N, V^N)$  each with the previous dynamics, where the  $N^{th}$  process follows IPD rate adaptation rule (11) with scale parameter  $\alpha = 1/N$  and gain parameter  $\kappa = g/N$ . Denoting by  $X_S^N(t)$  the corresponding process (defined via (16)), assume that the initial conditions admit limits

$$\lim_{N \rightarrow \infty} c_\ell^N(0) = c_\ell, \quad \lim_{N \rightarrow \infty} \frac{1}{N} X_S^N(0) = v_S$$

for all  $\ell \in \mathcal{L}$  and  $S \in \mathcal{T}_s$ . Then the rescaled processes

$$\begin{aligned} \hat{c}_\ell^N(t) &= c_\ell^N(Nt), \quad \ell \in \mathcal{L}, \\ Y_S^N(t) &:= \frac{1}{N} X_S^N(Nt), \quad S \in \mathcal{T}_s, \end{aligned}$$

converge in probability, for the topology of uniform convergence on finite intervals, as  $N \rightarrow \infty$ , to the set of fluid trajectories  $S(c, v)$ .

We now provide the key steps of the proofs, the details being omitted for brevity.

*Proof:* Fix some  $S \in \mathcal{T}_s$ . One can construct the process  $Y_S^N$  from unit rate Poisson processes  $P_0, P_{S,ij}, i \in S, j \in \bar{S}$  such that

$$Y_S^N(t) = Y_S^N(0) + \frac{1}{N} P_0(\lambda Nt) - \frac{1}{N} \sum_{i \in S, j \in \bar{S}} P_{S,ij}(N f_{S,ij}^N(t)),$$

where  $f_{S,ij}^N(t) := \int_0^t \hat{c}_{ij}^N(u) \left(1 - |\mathbb{F}|^{-N(Y_S^N(u) - Y_{S-i}^N(u))}\right) du$ . The uniform law of large numbers for Poisson processes enables to rewrite the previous equation as

$$Y_S^N(t) = Y_S^N(0) + \lambda t - \sum_{i \in S, j \in \bar{S}} f_{S,ij}^N(t) + \epsilon_S^N(t), \quad (24)$$

where the error term  $\epsilon_S^N(t)$  goes to zero with  $N$  uniformly in  $t$  on any finite interval  $[0, T]$ . Using boundedness of capacities  $c_{ij}$  suffices to establish that the sequence of processes  $(Y_S^N)$  is tight for the topology of uniform convergence on any finite interval  $[0, T]$  (detailed arguments would parallel closely those in [7], pp.7-10).

Moreover, for any  $(i, j) \in \mathcal{L}$ , it holds that

$$\hat{c}_{ij}^N(t) = \hat{c}_{ij}^N(0) + \int_0^t g \left[ v_{+i-j}^N(u) - p_{ij}(\hat{c}^N(u)) \right] du, \quad (25)$$

where  $v_{+i-j}^N(u) := Y_{\{j\}}^N(u) - Y_{\{i,j\}}^N(u)$ . Equation (25) together with continuity of  $p_{ij}$  entails that the sequence of processes  $\hat{c}^N$  is also tight. The limiting process  $v_S$  of  $Y_S^N$  verifies (22) as well as the supermodularity condition (23) by definition (16). The latter follows from the fact that the inequality  $\dim(V_{S \cap T}) \leq \dim(V_S \cap V_T)$  holds for any  $S, T \in \mathcal{T}_s$ . Finally, by Equations (24–25), for any subsequence along which all processes  $\hat{c}^N, Y^N$  converge, the limiting process  $v_S$  must further verify

$$v_S(t) = v_S + \lambda t - \sum_{i \in S, j \in \bar{S}} \int_0^t \hat{c}_{ij}(u) (1 - a_{S-i,S}(u)) du,$$

where the process  $a_{S-i,S}$  must satisfy (21), as a limit of process  $u \rightarrow (1/|\mathbb{F}|)^{N(Y_S^N(u) - Y_{S-i}^N(u))}$ . Taking limits in Equation (25), expression (19) for the derivative of the limiting process of rates  $c_{ij}$  directly follows. ■

### C. Stationarity of the Optimal State for IPD with RLC

We now establish that an optimal operating point is a fixed point of the fluid trajectories (20–21). To state the result, we will use the following notation. We let  $(c_{ij}^*, (i, j) \in \mathcal{L}, \mu_S^*, S \in \mathcal{T}_s)$  denote optimal primal and dual variables of the optimization problem (1–2). We also let

$$v_S^* := \sum_{T \subset S} \mu_T^*, \quad S \in \mathcal{T}_s. \quad (26)$$

We then have the following:

*Theorem 2:* Any fluid trajectory  $(c, v)$  satisfying conditions (20–21) and with initial conditions  $c(0) = c^*$  and  $v(0) = v^*$  is stationary, in that:

$$\begin{aligned} \frac{d}{dt} c_\ell(0) &= 0, \quad \ell \in \mathcal{L}, \\ \frac{d}{dt} v_S(0) &= 0, \quad S \in \mathcal{T}_s. \end{aligned} \quad (27)$$

Before we turn to its proof, let us comment on the meaning of the result. It indicates that the IPD with RLC system, when initialized with contents at nodes such that the corresponding quantities  $X_S$  as defined by (16) are close to  $\alpha^{-1} v_S^*$ , and with capacities  $c_{ij}$  close to the optimal values  $c_{ij}^*$ , will not deviate from this optimal operating point, at which network cost is minimal. We believe that a stronger result holds, namely that from any initial conditions, the dynamics of IPD with RLC (at the fluid level) converge to this optimal operating point. While a proof of this conjecture is currently missing, we provide a stability result in the next subsection, and experimental evidence in Section VI, which support the conjecture.

We now provide auxiliary results needed to establish Theorem 2.

*Lemma 1:* Let  $\tau$  be a tree rooted at  $s$ , spanning all nodes in  $\mathcal{V}$ . Let  $A_1, A_2$  be two cuts in  $\mathcal{T}_s$  that satisfy the following *criticality* property: there is a unique edge of  $\tau$  that crosses from  $A_i$  to  $\bar{A}_i, i = 1, 2$ . Then the cut  $A_1 \cup A_2$  is also critical, provided its complementary set  $\bar{A}_1 \cup \bar{A}_2$  is non-empty. The proof of this Lemma is available in the companion technical report [9].

*Corollary 1:* Consider capacities  $(c_\ell, \ell \in \mathcal{L})$  that are sufficient for streaming at rate  $\lambda$ , i.e. condition (2) holds. Assume further that they are *minimal*, in the sense that if there exist capacities  $c'_\ell$  that are also sufficient for streaming at rate  $\lambda$ , and that verify  $c'_\ell \leq c_\ell, \ell \in \mathcal{L}$ , then necessarily  $c_\ell \equiv c'_\ell$ .

Given any two cuts  $A_1, A_2 \in \mathcal{T}_s$  that are  $c$ -critical, i.e. inequality (2) holds with equality, then their union  $A_1 \cup A_2$  is  $c$ -critical whenever  $\bar{A}_1 \cup \bar{A}_2$  is non empty.

*Proof:* Edmonds's theorem [5] establishes the following result. The capacities  $(c_\ell, \ell \in \mathcal{L})$  are sufficient for streaming at rate  $\lambda$  if and only if there exists a collection of spanning trees  $\tau$ , and associated capacities  $\lambda_\tau > 0$ , such that these trees can be “packed” into link capacities  $c_\ell$ , i.e.

$$c_\ell \geq \sum_{\tau} \lambda_\tau \mathbf{1}_{\{\ell \in \tau\}}, \quad \ell \in \mathcal{L}. \quad (28)$$

It is then easy to see that the capacities  $c_\ell$  are both sufficient and minimal if inequalities (28) hold with equality.

Let  $A_1, A_2$  be critical cuts in  $\mathcal{T}_s$  such that  $\overline{A_1 \cup A_2}$  is non-empty. For any tree  $\tau$  with  $\lambda_\tau > 0$ , criticality of  $A_i$  entails that there is a single edge of  $\tau$  crossing from  $A_i$  to  $\overline{A_i}$ , for  $i = 1, 2$ . Therefore, in view of Lemma 1, there is also a single edge of  $\tau$  crossing from  $A_1 \cup A_2$  into its complement. Since this holds for all spanning trees  $\tau$  with  $\lambda_\tau > 0$ , it easily follows that  $A_1 \cup A_2$  is also critical. ■

The ingredients for the proof of Theorem 2 are now in place.

*Proof:* (of Theorem 2). Consider first the derivative of  $c_{ij}$ . In view of (19) and of the choice of initial conditions (26) and  $\vec{c}(0) = \vec{c}^*$ , it reads

$$\frac{d}{dt}c_{ij}(0) = g[v_{+i-j}^* - p_{ij}(\vec{c}^*)].$$

However,  $v_{+i-j}^*$  coincides with  $\sum_{S:i \in S, j \notin S} \mu_S^*$ . Thus the optimality of  $(\vec{c}^*, \mu^*)$  implies that the above time derivative equals zero.

Consider now the derivative of  $v_S$ . Let us first consider the case where the associated multiplier  $\mu_S^*$  is strictly positive. Then, by the complementary slackness condition (7), necessarily  $c^*(S, \overline{S}) = \lambda$ . In addition, for all  $i \in S$ ,  $v_S^* - v_{S-i}^* \geq \mu_S^* > 0$ . Thus, the corresponding term  $a_{S-i,S}$  in Equation (20) is necessarily zero, in view of (21). This establishes that the time derivative of  $v_S$  is indeed zero.

We then consider the case where the multiplier  $\mu_S^*$  equals zero. We shall proceed by induction on the size  $|S|$  of  $S \in \mathcal{T}_s$  to show that the time derivative of  $v_S$  must be zero. To start the induction, consider first the smallest cut in  $\mathcal{T}_s$ , that is  $S = \{s\}$ . Assume that  $\mu_{\{s\}}^* = 0$  (the other case being already treated). The derivative of  $v_{\{s\}}$  is necessarily non-negative, for otherwise, with initial condition  $v_{\{s\}}(0) = 0$ , process  $v_{\{s\}}$  would take on negative values, which contradicts the properties of the fluid limits. It cannot be strictly positive, for otherwise, for small enough  $t > 0$ , one would have  $v_{\{s\}}(t) > 0$ , and hence  $a_s(t) = 0$  in (21), in which case this derivative would read, for small enough  $t$ ,

$$\frac{d}{dt}v_{\{s\}}(t) = \lambda - \sum_{j \neq s} c_{sj}(t) = \lambda - \sum_{j \neq s} c_{sj}^* + o(t) \leq o(t),$$

a contradiction with the assumption that this derivative is positive at  $t = 0$ .

Consider now  $S \in \mathcal{T}_s \setminus \{s\}$ . The induction hypothesis is that for all  $T \subset S$ ,  $T \neq S$ ,  $\frac{d}{dt}v_T(0) = 0$ . Distinguish two cases. First assume that

$$\forall i \in S, \quad v_S^* - v_{S-i}^* > 0. \quad (29)$$

Thus necessarily, the terms  $a_{S-i,S}$  must be zero, for all  $i \in S$ . In addition, (29) entails that for all  $i \in S$ , there exists a subset  $T_i \subset S$  such that  $i \in T_i$  and  $\mu_{T_i}^* > 0$ . In turn, this implies that each such cut  $T_i$  must be  $c^*$ -critical, in the sense of Corollary 1. Thus, applying inductively the result of the corollary, it follows that  $\cup_{i \in S} T_i$  is also  $c^*$ -critical. However, since  $i \in T_i$ , the set  $\cup_{i \in S} T_i$  coincides with  $S$ , which is

then  $c^*$ -critical. Combined with the fact that  $a_{S-i,S} = 0$ , this implies the desired result that  $\frac{d}{dt}v_S(0) = 0$ .

The last case to consider is when for some  $i \in S$ , Condition (29) is violated, that is  $v_{S-i}^* = v_S^*$ . By the induction hypothesis, for all such  $i$ ,  $v_{S-i}(t) = v_{S-i}^* + o(t)$ . Thus the derivative of  $v_S$  at zero cannot be negative, for otherwise it would violate inequality (22):  $v_S \geq v_{S-i}$ . Assume then that it is positive. It then follows by the induction hypothesis that, for small enough  $t$ ,  $a_{S-i,S}(t)$  must be zero for all  $i \in S$ . Thus the derivative of  $v_S(t)$  for small  $t$  reads

$$\lambda - \sum_{i \in S, j \notin S} c_{ij}(t) = \lambda - c^*(S, \overline{S}) + o(t) \leq o(t),$$

a contradiction with the assumption that  $\frac{d}{dt}v_S(0)$  is positive. ■

#### D. Stability of IPD-RLC for Separable Costs

So far we established that at the fluid scale, IPD-RLC does not deviate from the optimal operating point once it has reached it. This begs the question of whether it is attracted to this optimal operating point, from arbitrary initial configurations. Lacking an answer at this stage, we consider instead the weaker property of convergence to a bounded set containing the optimal state. We establish that this indeed holds, for specific *separable* cost functions.

For this result we consider a different scaling than the one we presented in Subsection V-B. We suppose that the  $X(t)$  vary “slowly” and that the rates  $c_{ij}(t)$  are computed on a more rapid time scale. This would correspond to the fluid trajectories obtained in the limit of small scale parameters  $\alpha$ , but with fixed, non-vanishing gain parameters  $\kappa$  in (11).

We thus consider fluid trajectories as in Definition 1, with the exception that Equation (19) is replaced by the condition

$$[v_{+i-j}(t) - p_{ij}(\vec{c}(t))]_{c_{ij}(t)}^+ = 0, \quad (i, j) \in \mathcal{L}. \quad (30)$$

Furthermore, we assume that the cost function is separable, i.e. it can be explicitly written as

$$\Gamma(\vec{c}) = \sum_{\ell \in \mathcal{L}} \Gamma_\ell(c_\ell),$$

where the individual cost functions  $\Gamma_\ell$  are assumed convex and differentiable in the domain  $[0, \bar{c}_\ell)$ , with  $\bar{c}_\ell > 0$ , where they are finite. With such separable costs, the above condition (30) simplifies to

$$c_{ij}(t) = (\Gamma'_{ij})^{-1}(v_{+i-j}(t)). \quad (31)$$

We finally assume that streaming is feasible at rate  $\lambda + \epsilon$  for some positive  $\epsilon$ , which in the present context is equivalent to the following property:

There exist capacities  $\hat{c}_{ij}, (i, j) \in \mathcal{L}$  for which streaming at rate  $\lambda + \epsilon$  is feasible, and a positive value  $v_{\min} > 0$  such that

$$\forall \ell \in \mathcal{L}, \quad v \geq v_{\min} \Rightarrow (\Gamma'_\ell)^{-1}(v) \geq \hat{c}_\ell. \quad (32)$$

We are now in a position to establish the following

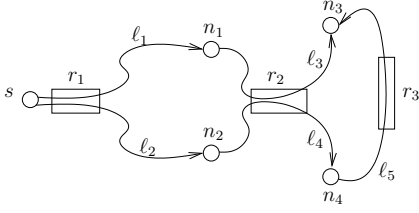


Fig. 2: A network with shared physical links.

**Theorem 3:** Consider separable cost functions as above. Then any fluid trajectories with fast rate adaptation as described by (31) are asymptotically bounded, i.e. there exists a constant  $A > 0$  such that the quantities  $v_S(t)$  verify  $\limsup_{t \rightarrow \infty} v_S(t) \leq A$ .

Before proving the Theorem, we need the following

**Lemma 2:** Let  $r \in (0, 1)$  and  $v \in \mathbb{R}_+^{|\mathcal{T}_s|}$  such that the supermodularity property (23) is verified. Denoting  $v_{+i-j} = v_{\{j\}} - v_{\{i,j\}}$ , then for all  $S \in \mathcal{T}_s$ , all  $i \in S$ ,  $j \in \bar{S}$ , the following inequality holds:

$$v_{+i-j} < rv_S \Rightarrow (1-r)^{|S|-1} v_{S-i} > (1-r)^{|S|} v_S. \quad (33)$$

*Proof:* Note that by definition of  $v_{+i-j}$  and by writing (23) for cuts  $S$  and  $\{i, j\}$ , one has  $v_S \leq v_{S-i} + v_{+i-j}$ . Using the assumption  $v_{+i-j} < rv_S$ , it follows that  $v_S < v_{S-i} + rv_S$ , and the conclusion easily follows. ■

*Proof:* (of Theorem 3). Fix some  $r \in (0, 1)$ . Consider the following Lyapunov function:

$$L(v) = \sup_{S \in \mathcal{T}_s} (1-r)^{|S|} v_S.$$

If we denote by  $S^*$  the cut in  $\mathcal{T}_s$  achieving the maximum in the definition of  $L$  at a certain point in time, then for all  $i \in S^*$ ,  $j \notin S^*$ , we have  $v_{+i-j} \geq rv_{S^*}$ . Indeed, if this did not hold, by Lemma 2 the cut  $S' = S^* - i$  would satisfy  $(1-r)^{|S'|} v_{S'} > (1-r)^{|S^*|} v_{S^*}$ , a contradiction of  $S^*$ 's maximality.

Moreover, for all  $i \in S^*$ , necessarily  $v_{S^*} \geq (1-r)^{-1} v_{S^*-i} > v_{S^*-i}$ , since  $r < 1$ . Thus the quantities  $a_{S^*-i, S^*}$  in the derivative (20) are zero. Finally, whenever  $rv_{S^*}$  is larger than  $v_{\min}$ , necessarily  $v_{+i-j} \geq v_{\min}$ , and thus by the stability assumption (32), it follows that  $\frac{d}{dt} v_{S^*} \leq \lambda - \sum_{i \in S^*, j \notin S^*} \hat{c}_{ij} \leq -\epsilon$ . We have thus established that  $L(v)$  strictly decreases whenever  $rv_{S^*} \geq v_{\min}$ . A fortiori, it decreases whenever  $L(v) \geq r^{-1} v_{\min}$ . It therefore follows that asymptotically,  $L(v)$  must be below  $r^{-1} v_{\min}$ . Thus for each  $S \in \mathcal{T}_s$ , one has  $\limsup_{t \rightarrow \infty} (1-r)^{|S|} v_S(t) \leq r^{-1} v_{\min}$ , and the result of the Theorem follows by taking the constant  $A$  equal to  $r^{-1} (1-r)^{-|\mathcal{V}|} v_{\min}$ . ■

## VI. NUMERICAL EVALUATIONS

In this section we present numerical evaluations for IPD and we describe the influence of various implementation parameters on the algorithm's performance. For our simulations we consider the network depicted in Figure 2. For this network shared physical links  $r_1, r_2$  and  $r_3$  have a maximum physical rate of 11. The source  $s$  generates content at rate  $\lambda = 5$ . We wish to minimize congestion. We use

a cost function which sets overlay link prices to the sum of quantities inversely proportional to the spare capacities on the physical links used. This price function (3) has been discussed in Section III. The optimal rate allocation in this case is  $c_{s1}^* = c_{s2}^* = c_{24}^* = c_{43}^* = 5$  and  $c_{13}^* = 0$ . Node 3 can receive all of its content from node 4. Any other strategy would have a higher global cost, due to the strict convexity of the cost function. Intuitively, instead of imposing a high load on physical link  $r_2$ , it is globally better to make use of the free resource  $r_3$ .

We test IPD under the following scenarios:

a) *Idealized Coding:* We simulate explicitly the evolution of quantities  $X_S(t)$  described in Section V. These simulations are a best case scenario for the achievable performance of Random Linear Coding. The simulation is feasible, since there are only  $2^{5-1} = 16$  variables associated with the backlogs. It is an idealized case in the sense that we use an infinite coding field size. Thus the only limitation is the “content bottleneck” (i.e. for every link  $(i, j)$  such that  $X_{+i-j} > 0$  and every  $S \in \mathcal{T}_s$  such that  $i|_S$ , we consider  $r(S, S-i) = \mathbf{1}_{\{X_S = X_{S-i}\}}$  instead of (18) from Section V). Furthermore, nodes need to receive all the data from the destination to be able to decode the content (and start playing it).

b) *Generation Coding:* We use a discrete packet-level simulator for which we implement Random Linear Coding using generations of a fixed size  $K$ . In this setting nodes need to receive  $K$  linearly independent combinations within a generation to be able to decode it and start playing. For all links  $(i, j)$ , nodes  $i$  select a “useful” generation for  $j$  at random, compute a random linear combination within this generation and then forward it to node  $j$ . This is a realistic setting. We consider several generation sizes, as well as two finite fields:  $\mathbb{F}_2$  and  $\mathbb{F}_{2^8}$  for evaluating the performance of our scheme.

c) *Random Useful:* Finally, we implement Random Useful in our packet-level simulator.

In our simulations receivers become active at  $t = 300$  time units. By this time, the source generates about 1500 packets, a delay from which the receivers will want to recover. We simulate several scenarios with three generation sizes and two finite fields used. Define the Relative Mean Error (RME) of an execution as the relative error of the mean global cost observed  $\bar{\gamma}$  with respect to the optimal value  $\gamma^*$ , namely  $\frac{\bar{\gamma} - \gamma^*}{\gamma^*}$ .

In Figure 3 we consider the Idealized Coding case for two values of  $\alpha$  (we used  $\kappa = 0.06$ ). We plot the receiver delays with respect to the source ( $X_{+s-j}$ ) and the rates at which overlay links are being used for the two cases. The curve converging to the lowest value corresponds to node  $n_3$  for the delay plots and to link  $\ell_3 = (1, 3)$  for the rate plots. The cost function we are considering is not separable (as it was assumed in the proof of Theorem 3). Nevertheless, we notice that our system stabilizes after roughly 5,000 time units. This observation reinforces our conjecture that IPD is stable for more general cost functions. Furthermore, we observe that the instance using a lower  $\alpha$  is closer to the optimum in two ways: First, link  $\ell_3$  is less utilized than in the case of the instance



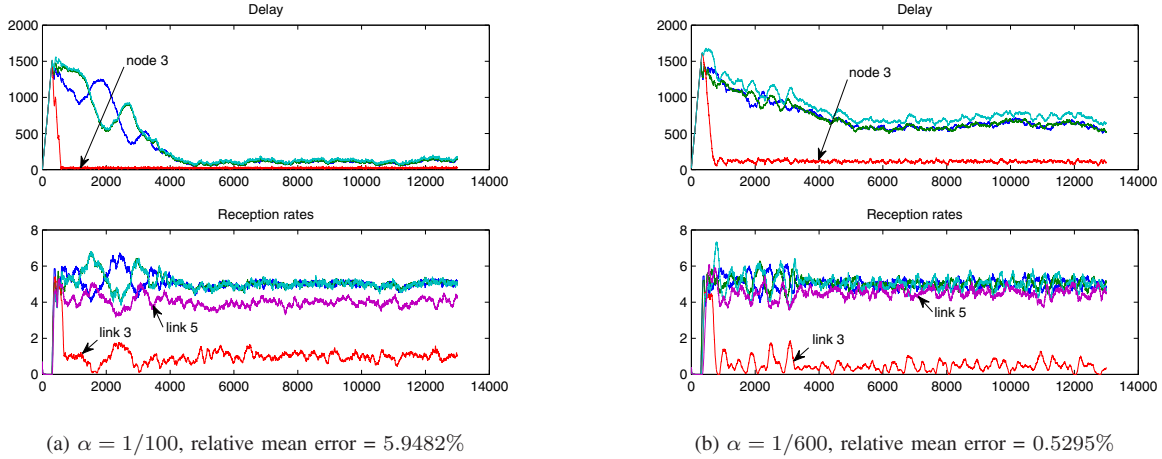


Fig. 3: Influence of  $\alpha$  – delay of node  $n_3$  and rate usage of links  $\ell_3 = (1, 3)$  and  $\ell_5 = (4, 3)$  are indicated.

using a higher  $\alpha$ . Second, the relative error is almost 10 times smaller in this case. However, the price we pay for approaching the optimum is a difference in delays at receiver nodes of almost an order of magnitude between the two instances. Since in a real implementation low delays are desirable, for the next part of this Section we choose the higher value of  $\alpha = 0.01$ .

Let us now observe performance of the Generation Coding scheme versus the Random Useful scheme. As previously stated, we use  $\alpha = 0.01$  and  $\kappa = 0.06$  in (11). In Table I we

Scheme	Field size	Generation size	RME
Idealized Coding	–	–	5.95%
Generation Coding	256	150	11.01%
		100	10.16%
		50	12.01%
	2	150	18.33%
		100	21.43%
		50	40.91%
Random Useful	–	–	12.94%

TABLE I: Relative Mean Error

give the Relative Mean Error of an execution instance for each of the cases, for a time span of 20,000 time units. We notice that in this setting, when ruling out the Idealized Coding case, the best performance is obtained by performing coding over  $\mathbb{F}_{2^8}$ . The generation size parameter has the anticipated impact on performance: the higher the generation size, the closer we are to the optimum. It is surprising that IPD in a Random Useful setting, while proven suboptimal in Section IV, gives similar performance to IPD in a Generation Coding setting for  $\mathbb{F}_{2^8}$  and outperforms it when the field size is equal to 2. Indeed, while for  $\mathbb{F}_{2^8}$  virtually all transfers are linearly independent combinations, for  $\mathbb{F}_2$  the number of transferred linearly dependent combinations becomes non-negligible.

## VII. CONCLUSION

In this paper we proposed a fully distributed cost-efficient rate control scheme for live streaming peer-to-peer systems.

We formulated our goal as an optimization problem having a convex cost function. Using the backlog sizes as approximations for the dual variables we introduced *Implicit-Primal-Dual*, our rate control scheme. We managed to show the advantage of using Network Coding over a classic Random Useful approach. Namely, while the two have the same feasible capacity region in the considered setting, we proved that for Random Linear Coding IPD has a fixed point at the global optimum while this does not hold in the Random Useful setting. Our numerical evaluations showed that IPD still performs reasonably well for Random Useful packet networks.

## REFERENCES

- [1] Minghua Chen, Miroslav Ponec, Sudipta Sengupta, Jin Li, and Philip A. Chou. Utility maximization in peer-to-peer systems. In *SIGMETRICS '08: Proceedings of the 2008 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 169–180, New York, NY, USA, 2008. ACM.
- [2] Desmond S. Lun, Niranjan Ratnakar, Muriel Médard, Ralf Koetter, David R. Karger, Tracey Ho, Ebad Ahmed, and Fang Zhao. Minimum-cost multicast over coded packet networks. *IEEE/ACM Trans. Netw.*, 14(SI):2608–2623, 2006.
- [3] S. Bhadra, S. Shakkottai, and P. Gupta. Min-cost selfish multicast with network coding. *Information Theory, IEEE Transactions on*, 52(11):5077–5087, 2006.
- [4] Raymond W. Yeung, Shuo-Yen Robert Li, Ning Cai, and Zhen Zhang. Network coding theory: single sources. *Commun. Inf. Theory*, 2(4):241–329, 2005.
- [5] Jack Edmonds. Edge-disjoint branchings. *R. Rustin (Ed.), Combinatorial Algorithms, Algorithmics press*, pages 21–31, 1972.
- [6] Rayadurgam Srikant. *The Mathematics of Internet Congestion Control (Systems and Control: Foundations and Applications)*. SpringerVerlag, 2004.
- [7] Laurent Massoulié and Andrew Twigg. Rate-optimal schemes for peer-to-peer live streaming. *Perform. Eval.*, 65(11-12):804–822, 2008.
- [8] Siddhartha Annapureddy, Saikat Guha, Christos Gkantsidis, Dinan Gunawardena, and Pablo Rodriguez. Exploring VoD in P2P Swarming Systems. In *Proceedings of the 26th Annual IEEE Conference on Computer Communications Mini Symposiums (INFOCOM '07 Mini Symposium)*, Anchorage, Alaska, May 2007.
- [9] Dan-Cristian Tomozei and Laurent Massoulié. Flow control for cost-efficient peer-to-peer streaming. Technical report, Thomson, 2009. Available as <http://www.thlab.net/~dantom/pub/nc-rate.pdf>.